Quantum basin hopping with gradient-based local optimisation

David Bulger*

February 1, 2008

Abstract

The quantum basin hopping algorithm for continuous global optimisation combines a local search with Grover's algorithm, and can locate the global optimum using effort proportional to the square root of the number of basins. This article establishes that Jordan's quantum gradient estimation method can be incorporated into the quantum basin hopper, providing an extra acceleration proportional to the domain dimension.

MSC2000 Subject Classification: 90C30, 68Q99, 68Q25
Key words and phrases: global optimisation, conjugate gradient, quantum computation, Grover's algorithm, quantum basin hopper.

1 Introduction

Recent articles [1, 4, 7, 8, 11] have studied the application of Grover's quantum algorithm [2, 9] to the general unconstrained global optimisation problem. Given a domain point and corresponding objective function value, we can dynamically construct a quantum oracle classifying any other domain point as either better or not. Using this oracle in Grover's algorithm produces an implementation of pure adaptive search [13], that is, an improving sequence of domain points, each uniformly distributed among all points better than the previous. Details in [1] show that the effort required by this approach is on the order of the square root of the effort required by pure random search [3].

Of course, for many problems, more sophisticated methods than pure random search can be used on a Turing computer. Typically in these methods the idea is to be "guided" toward optimal domain points by the objective function's "terrain", that is, the shape of its graph. For instance, the multistart method is a simple hybrid of pure random search and a local descent search, in which independent exhaustive local searches are conducted from many randomly chosen initial points. The quantum basin hopper [5], a quantum analogue of the multistart method, combines the quadratic acceleration due to Grover's algorithm with the potentially huge acceleration over pure random search to be gained by using a local search.

In practice, objective functions in continuous domains are usually differentiable, at least piecewise. Being guided by the terrain, in this case, will largely mean responding to the gradient, an analytic expression for which may or may not be available. Local optimisation techniques for such problems generally fall into two categories. Gradient methods, such as gradient descent and conjugate gradient, require explicit computation of the gradient at each sample point. On the other hand, gradient-free methods, such as the Nelder-Meade simplex method and simulated annealing, attempt to respond to the gradient in an indirect way, without requiring its explicit computation. (The methods named are all described in [12].) Much of the reason for gradient-free methods is that, when an analytic expression is available for the objective function but not for its derivative, the number of function evaluations required near a point to estimate the gradient increases linearly with the domain dimension. In many problems of practical interest, the domain dimension is large enough to make this unpractical.

---
*Lecturer, Department of Statistics, Macquarie University, NSW 2109, Australia, dbulger@efs.mq.edu.au.

The *quantum* complexity of estimating an objective function's gradient, based solely on function evaluations, is constant, rather than linear, in dimension. Specifically, [10] shows that the gradient can be estimated at a point using just one (superposed) evaluation of the objective function. In [6] it is established, under mild regularity conditions, that the error in the gradient estimate can be reduced to any required level, by appropriately selecting the numerical encoding and calculation precision and other parameters of the gradient estimation algorithm.

Therefore it seems that gradient information is cheap in quantum computation, and ought to be incorporated into local optimisation methods, and into the quantum basin hopper. This paper discusses that possibility. The difficulty is that, while [5] assumes the local search to act deterministically, quantum gradient estimation produces in general a nontrivial superposition of computational basis states. This is not simply a matter of numerical error; the indeterminacy, together with the marking oracle used in the Grover search, causes unwanted entanglement between the search register and the gradient estimation register. We will argue that the error bound established in [6] implies that the algorithm's performance is substantially unaffected by this entanglement.

This paper's structure is as follows. Section 2 presents the required details of the quantum basin hopper. In Section 3, we consider the error introduced by the indeterminacy of quantum gradient estimation, and conceptually divide it into two different kinds of error, "large" and "small". Section 4 shows that the amplitude of the "large" errors can be made too small to disrupt the quantum basin hopper. Section 5 shows that the numerical scale of the "small" errors adequately restricts the effect of the entanglement. Section 6 concludes the paper and speculates on related future research.

2   The quantum basin hopper

The quantum basin hopper algorithm is described in [5]. The specifically quantum part of the algorithm, denoted $\mathrm{GBS}(Y, r, L)$ below, is nested within two loops which can be executed classically.

Quantum basin hopper with parameters (λ, M, L)
    randomly choose $x \in S$
    descend from x to a local minimum x_{cand}
    $y_{cand} \leftarrow f(x)$
    repeat
        $x \leftarrow x_{cand}$
        $y \leftarrow y_{cand}$
        $m \leftarrow 1$
        repeat
            choose an integer r uniformly at random from $\{0, \ldots, \lceil m - 1 \rceil\}$
            set x_{cand} and y_{cand} by measuring the output of $\mathrm{GBS}(Y, r, L)$
            $m \leftarrow \lambda m$
        until $y_{cand} < Y$ **or** $m > M$
    until $y_{cand} \geq Y$

The quantum part $\mathrm{GBS}(Y, r, L)$ simply uses Grover's algorithm with r rotations to seek a domain point whose containing basin extends below Y. (Basins are exlored using L iterations of the local search.) $\mathrm{GBS}(Y, r, L)$ operates on a Hilbert space $\mathcal{D}^{\otimes(L+1)} \otimes \mathcal{R}$, encoding the state of a system containing $L+1$ registers storing points in the domain of f, and one register storing a value of f. It proceeds as follows. The state is initialised to $|s\rangle \otimes |0_{\mathcal{D}}\rangle^{\otimes L} \otimes |0_{\mathcal{R}}\rangle$, where $|s\rangle$ is the equal-amplitude state in $\mathcal{D}$. Then, the operator

$$U_f U_{d,L} \cdots U_{d,1} (U_s U_{d,1}^{-1} \cdots U_{d,L}^{-1} U_< U_f U_{d,L} \cdots U_{d,1})^r \tag{1}$$

is applied. Finally, the last two registers are measured, and their values are returned as x_{cand} and y_{cand}. The factors process data as follows:

$$U_f |x_0\rangle \otimes \cdots \otimes |x_L\rangle \otimes |y\rangle \;\; = \;\; |x_0\rangle \otimes \cdots \otimes |x_L\rangle \otimes |y + f(x_L)\rangle,$$

$$U_s|x_0\rangle \otimes \cdots \otimes |x_L\rangle \otimes |y\rangle = ((2|s\rangle\langle s| - I)|x_0\rangle) \otimes \cdots \otimes |x_L\rangle \otimes |y\rangle,$$

$$U_<|x_0\rangle \otimes \cdots \otimes |x_L\rangle \otimes |y\rangle = \begin{cases} |x_0\rangle \otimes \cdots \otimes |x_L\rangle \otimes |y\rangle, & y \geq Y, \\ -|x_0\rangle \otimes \cdots \otimes |x_L\rangle \otimes |y\rangle, & y < Y, \end{cases}$$

(where, for unitarity, the addition $y + f(x_L)$ is modulo the ordinate register's storage capacity). The operator $U_{d,k}$ updates the $(k+1)$th register on the basis of the kth, representing a single iteration of the local minimisation. In [5], it is assumed to act deterministically, so that $U_{d,L} \cdots U_{d,1}$ maps the computational basis state

$$|x_0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle \otimes |0\rangle$$

to the computational basis state

$$|x_0\rangle \otimes |x_1\rangle \otimes \cdots \otimes |x_L\rangle \otimes |0\rangle,$$

where $x_1, \ldots, x_L$ is the local search's sample path when started at x_0.

3 Errors due to quantum indeterminacy in gradient estimates

In this article though, we assume instead that each operation $U_{d,k}$ involves a single quantum gradient estimation, so that the operation (1) involves $(2r+1)L$ of them. The effect of the indeterminacy inherent in the gradient estimation method on Grover's algorithm requires contemplation. The terminus of the local minimisation sample path resulting from each domain point will be a quantum superposition, peaked around a point near the bottom of the point's containing basin, but having a nonzero amplitude at every domain point in general. Accordingly, the marking oracle involved in Grover's algorithm, which is here intended to identify points in basins extending below Y, will always slightly entangle the register storing the sample path's origin with the circuitry estimating the gradient.

The aim of this article is to demostrate that Grover's algorithm, and therefore the quantum basin hopper, can still function despite this entangling. This will be done by considering separately the effects of "large" errors and "small" errors in the gradient estimates. The error in a gradient estimate is simply the magnitude of the difference between the estimate and the true gradient, and it is considered "large" if it exceeds δ, and "small" otherwise, where δ is a positive constant of our choosing.

4 Large errors

If $|g_n\rangle$ is the nth gradient estimate, let $|g_n'\rangle$ denote the normalised projection of $|g_n\rangle$ onto the span of estimates with error no larger than δ. Theorem 1 of [6] implies that, for any positive ϵ and δ, sufficiently careful quantum gradient estimation will ensure that $\||g_n'\rangle - |g_n\rangle\| < \epsilon$. Let $|\phi\rangle$ represent the computer's state just before measurement at the end of GBS(Y, r, L), and let $|\phi'\rangle$ represent the corresponding state if the gradient estimates $|g_n\rangle$ are replaced with $|g_n'\rangle$. Clearly $\||\phi'\rangle - |\phi\rangle\| < (2r+1)L\epsilon$, which can be made as small as required. Therefore we can safely ignore large errors.

5 Small errors

It remains only to show that the quantum basin hopper can tolerate small errors, that is, that if gradient estimate errors were uniformly bounded by a sufficiently small δ, the quantum calculation GBS(Y, r, L) would locate a basin extending lower than Y with a high probability (for suitable rotation count r, and assuming such basins exist).

Let $f : D \to \mathbb{R}$ be the objective function before discretisation. Suppose $D \subset \mathbb{R}^p$, and let $(D, \mathcal{F}, \mu)$ be a measure space with $\mu(D)$ finite and f measurable.

Let $d_L(x_0, u_1, \ldots, u_L)$ denote the Lth iterate of the local search, starting from x_0, and assuming that the kth gradient estimate includes an error of u_k. Let

$$
\begin{aligned}
W_+ &= \{x : d_L(x, 0, \ldots, 0) > Y\}, \\
W_Y &= \{x : d_L(x, 0, \ldots, 0) = Y\}, \text{ and} \\
W_- &= \{x : d_L(x, 0, \ldots, 0) < Y\}.
\end{aligned}
$$

For almost all Y, W_Y will have measure zero. Ordinarily, for a wide class of local optimisation methods, d_L is continuous in all its arguments, almost everywhere. Thus, letting B_δ represent the closed ball around $0 \in \mathbb{R}^p$ with radius δ, the set

$$
W_{Y,\delta} = \{x : d_L(x, B_\delta, \ldots, B_\delta) \text{ contains elements less than and greater than } Y\}
$$

decreases, as $\delta \to 0$, to W_Y. This means that we can make $W_{Y,\delta}$ as small as we like by choosing δ small enough, that is, by computing the gradient estimates with sufficient precision.

Now suppose S discretises D, so that $\mathcal{D}$ is spanned by $\{|x\rangle : x \in S\}$. Partition S into S_α, S_β and S_γ, where

$$
\begin{aligned}
S_\beta &= S \cap W_{Y,\delta}, \\
S_\alpha &= (S \backslash S_\beta) \cap W_+, \\
S_\gamma &= (S \backslash S_\beta) \cap W_-.
\end{aligned}
$$

Section 2 indicates that the quantum search involves $L + 1$ domain-point registers and one range value register; in practice, there will be extra storage for "scratch space", including the register for the gradient estimation sampling grid. Taking all of this circuitry in account, define the Hilbert space $\mathcal{O}$ such that $\mathcal{D} \otimes \mathcal{O}$ represents the state all registers involved in the quantum search, with $\mathcal{D}$ representing the register storing x_0, and $\mathcal{O}$ representing all other registers. Let $|0_\mathcal{O}\rangle$ represent the initialised state of $\mathcal{O}$.

Let

$$
\begin{aligned}
N_\alpha = |S_\alpha|, \qquad |\alpha\rangle &= \sum \{|x\rangle \otimes |0_\mathcal{O}\rangle : x \in S_\alpha\}/\sqrt{N_\alpha}, \\
N_\beta = |S_\beta|, \qquad |\beta\rangle &= \sum \{|x\rangle \otimes |0_\mathcal{O}\rangle : x \in S_\beta\}/\sqrt{N_\beta}, \\
N_\gamma = |S_\gamma|, \qquad |\gamma\rangle &= \sum \{|x\rangle \otimes |0_\mathcal{O}\rangle : x \in S_\gamma\}/\sqrt{N_\gamma}.
\end{aligned}
$$

Let

$$
G = U_s U_{d,1}^{-1} \cdots U_{d,L}^{-1} U_< U_f U_{d,L} \cdots U_{d,1},
$$

so that GBS(Y, r, L) simply applies $U_f U_{d,L} \cdots U_{d,1} G^r$ to $|s\rangle \otimes |0_\mathcal{O}\rangle$.

Then

$$
\begin{aligned}
G|\alpha\rangle &= \frac{N_\alpha - N_\beta - N_\gamma}{N}|\alpha\rangle + \frac{2\sqrt{N_\alpha N_\beta}}{N}|\beta\rangle + \frac{2\sqrt{N_\alpha N_\gamma}}{N}|\gamma\rangle \text{ and} \\
G|\gamma\rangle &= -\frac{2\sqrt{N_\alpha N_\gamma}}{N}|\alpha\rangle - \frac{2\sqrt{N_\beta N_\gamma}}{N}|\beta\rangle + \frac{N_\alpha + N_\beta - N_\gamma}{N}|\gamma\rangle;
\end{aligned}
$$

however, $G|\beta\rangle$ is not a tensor multiple of $|0_\mathcal{O}\rangle$, and represents an entangled state. Let $|\beta_k\rangle$ denote $G^k|\beta\rangle$.

Define $\theta, \eta, \zeta \in [0, \pi/2]$ by

$$
\cos^2 \theta = \frac{N_\beta^2}{4N_\alpha N_\gamma} \qquad\qquad \cos^2 \eta = \frac{(N_\alpha - N_\gamma)^2}{(N_\alpha + N_\gamma)^2 - N_\beta^2} \qquad\qquad \cos^2 \zeta = \frac{(N_\alpha + N_\beta/2)^2}{NN_\alpha}.
$$

Define

$$
|\chi_+\rangle = \frac{1}{\sqrt{2}}\left(e^{i\theta}|\alpha\rangle - |\gamma\rangle\right) \qquad\qquad |\chi_-\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\theta}|\alpha\rangle - |\gamma\rangle\right).
$$

4

The Grover operation (not exactly a rotation in this context) acts on $|\chi_\pm\rangle$ as follows:

$$
\begin{aligned}
G|\chi_+\rangle &= \frac{1}{\sqrt{2}} U_s U_{d,1}^{-1} \cdots U_{d,L}^{-1} U_< U_f U_{d,L} \cdots U_{d,1} \left(e^{i\theta}|\alpha\rangle - |\gamma\rangle\right) \\
&= \frac{1}{\sqrt{2}} U_s \left(e^{i\theta}|\alpha\rangle + |\gamma\rangle\right) \\
&= \frac{|\alpha\rangle}{\sqrt{2}} \left(\frac{2e^{i\theta}N_\alpha}{N} + \frac{2\sqrt{N_\alpha N_\gamma}}{N} - e^{i\theta}\right) + \frac{|\gamma\rangle}{\sqrt{2}} \left(\frac{2e^{i\theta}\sqrt{N_\alpha N_\gamma}}{N} + \frac{2N_\gamma}{N} - 1\right) \\
&\quad + \frac{\sqrt{2N_\beta}}{N}\left(e^{i\theta}\sqrt{N_\alpha} + \sqrt{N_\gamma}\right)|\beta\rangle \\
&= \left(1 - \frac{2\sqrt{N_\gamma}}{N}\left(\sqrt{N_\gamma} + e^{i\theta}\sqrt{N_\alpha}\right)\right)|\chi_+\rangle + \frac{\sqrt{2N_\beta}}{N}\left(e^{i\theta}\sqrt{N_\alpha} + \sqrt{N_\gamma}\right)|\beta\rangle \\
&= re^{-i\eta}|\chi_+\rangle + \frac{\sqrt{2N_\beta}}{N}\left(e^{i\theta}\sqrt{N_\alpha} + \sqrt{N_\gamma}\right)|\beta\rangle, \quad\text{where} \\
r &= \sqrt{\frac{N_\alpha - N_\beta + N_\gamma}{N}}, \quad\text{and similarly} \\
G|\chi_-\rangle &= re^{i\eta}|\chi_-\rangle + \frac{\sqrt{2N_\beta}}{N}\left(e^{-i\theta}\sqrt{N_\alpha} + \sqrt{N_\gamma}\right)|\beta\rangle.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
G^k|\chi_+\rangle &= r^k e^{-ik\eta}|\chi_+\rangle + \frac{\sqrt{2N_\beta}}{N}\left(e^{i\theta}\sqrt{N_\alpha} + \sqrt{N_\gamma}\right)\sum_{j=0}^{k-1} r^j e^{-ij\eta}|\beta_{k-1-j}\rangle \quad\text{and} \\
G^k|\chi_-\rangle &= r^k e^{ik\eta}|\chi_-\rangle + \frac{\sqrt{2N_\beta}}{N}\left(e^{-i\theta}\sqrt{N_\alpha} + \sqrt{N_\gamma}\right)\sum_{j=0}^{k-1} r^j e^{ij\eta}|\beta_{k-1-j}\rangle.
\end{aligned}
$$

Noting that the initial state $|s\rangle \otimes |0_\mathcal{O}\rangle$ equals

$$
-\frac{i}{\sqrt{2N}\sin\theta}(\sqrt{N_\alpha} + e^{-i\theta}\sqrt{N_\gamma})|\chi_+\rangle + \frac{i}{\sqrt{2N}\sin\theta}(\sqrt{N_\alpha} + e^{i\theta}\sqrt{N_\gamma})|\chi_-\rangle + \sqrt{\frac{N_\beta}{N}}|\beta\rangle,
$$

we have

$$
\begin{aligned}
G^k|s\rangle \otimes |0_\mathcal{O}\rangle &= \frac{ir^k}{\sqrt{2N}\sin\theta}\left(-e^{-ik\eta}(\sqrt{N_\alpha} + e^{-i\theta}\sqrt{N_\gamma})|\chi_+\rangle + e^{ik\eta}(\sqrt{N_\alpha} + e^{i\theta}\sqrt{N_\gamma})|\chi_-\rangle\right) + \sqrt{\frac{N_\beta}{N}}|\beta_k\rangle \\
&\quad + \frac{i\sqrt{N_\beta}}{N\sqrt{N}\sin\theta}\sum_{j=0}^{k-1} r^j\left((\sqrt{N_\alpha} + e^{i\theta}\sqrt{N_\gamma})^2 e^{i(j\eta-\theta)} - (\sqrt{N_\alpha} + e^{-i\theta}\sqrt{N_\gamma})^2 e^{i(\theta-j\eta)}\right)|\beta_{k-1-j}\rangle \\
&= \frac{r^k}{\sqrt{N}\sin\theta}\left((\sqrt{N_\alpha}\sin(k\eta) + \sqrt{N_\gamma}\sin(k\eta+\theta))|\gamma\rangle - (\sqrt{N_\alpha}\sin(k\eta-\theta) + \sqrt{N_\gamma}\sin(k\eta))|\alpha\rangle\right) \\
&\quad + \sqrt{\frac{N_\beta}{N}}|\beta_k\rangle \\
&\quad + \frac{i\sqrt{N_\beta}}{N\sqrt{N}\sin\theta}\sum_{j=0}^{k-1} r^j\left((\sqrt{N_\alpha} + e^{i\theta}\sqrt{N_\gamma})^2 e^{i(j\eta-\theta)} - (\sqrt{N_\alpha} + e^{-i\theta}\sqrt{N_\gamma})^2 e^{i(\theta-j\eta)}\right)|\beta_{k-1-j}\rangle \\
&= \frac{r^k}{\sin\theta}\left(\sin(k\eta+\zeta)|\gamma\rangle - \sin(k\eta+\zeta-\theta)|\alpha\rangle\right) \\
&\quad + \sqrt{\frac{N_\beta}{N}}\left(|\beta_k\rangle + \frac{2}{\sin\theta}\sum_{j=0}^{k-1} r^j \sin(\theta - 2\zeta - j\eta)|\beta_{k-1-j}\rangle\right).
\end{aligned}
$$

In particular, noting that $|\gamma\rangle$ is orthogonal to $|\alpha\rangle$ and $|\beta\rangle$ but not necessarily to $|\beta_k\rangle$ for

$k > 0$, we have

$$\langle \gamma | G^k | s \rangle \geq \frac{r^k \sin(k\eta + \zeta)}{\sin\theta} - \frac{2}{\sin\theta}\sqrt{\frac{N_\beta}{N}}\sum_{j=0}^{k-1} r^j |\sin(\theta - 2\zeta - j\eta)| \tag{2}$$

$$= \frac{r^k \sin(k\eta + \zeta)}{\sin\theta} - \frac{2}{\sin\theta}\sqrt{\frac{N_\beta}{N}}\sum_{j=0}^{k-1} r^j \sin(\theta - 2\zeta - j\eta) \text{ for small } k$$

$$= \frac{r^k \sin(k\eta + \zeta)}{\sin\theta} - \frac{i}{\sin\theta}\sqrt{\frac{N_\beta}{N}}\left(e^{i(2\zeta-\theta)}\frac{1 - r^k e^{ik\eta}}{1 - re^{i\eta}} - e^{-i(2\zeta-\theta)}\frac{1 - r^k e^{-ik\eta}}{1 - re^{-i\eta}}\right).$$

This bound is maximised when

$$\frac{2\sqrt{N_\beta}}{\sqrt{N}}\Re\left(e^{i(2\zeta-\theta+k\eta)}\frac{\eta - i\ln r}{1 - re^{i\eta}}\right) = \ln r \sin(k\eta + \zeta) + \eta\cos(k\eta + \zeta),$$

that is, when

$$k = \frac{1}{\eta}\operatorname{atn}\frac{2\sqrt{N_\beta/N}\rho\cos(2\zeta - \theta + \phi) - \ln r \sin\zeta - \eta\cos\zeta}{2\sqrt{N_\beta/N}\rho\sin(2\zeta - \theta + \phi) + \ln r \cos\zeta - \eta\sin\zeta},$$

where

$$\rho e^{i\phi} = \frac{\eta - i\ln r}{1 - re^{i\eta}}.$$

This reduces to the familiar situation when $N_\beta = 0$; note that then we have

$$\theta = \frac{\pi}{2}, \qquad \cos\eta = 1 - 2\frac{N_\gamma}{N}, \qquad \zeta = \eta/2, \qquad r = 1,$$

and the optimal k is $\pi/2\eta - 1/2$.

More tractibly, from (2) we have

$$\langle \gamma | G^k | s \rangle \geq \frac{1}{\sin\theta}\left(r^k \sin(k\eta + \zeta) - 2k\sqrt{N_\beta/N}\right).$$

If $N_\beta \ll N_\gamma \ll N_\alpha$, then the usual estimate $k = (\pi/4)\sqrt{N/N_\gamma}$ gives

$$\frac{1}{\sin\theta}\left(r^k \sin(k\eta + \zeta) - 2k\sqrt{N_\beta/N}\right) \approx 1 - \frac{\pi}{4}\sqrt{\frac{N_\beta}{N_\gamma}}\left(1 + \sqrt{\frac{N_\beta}{N}}\right);$$

thus if the number of iterations is chosen optimally, the error rate is on the order of $\sqrt{N_\beta/N_\gamma}$. This will usually be overshadowed by errors due to unknown target size, which are already known to be manageable [2].

To summarise, "small" errors complicate the required implementations of Grover's algorithm by introducing an entangling subset of the search space. In the usual search algorithm, the search space is partitioned into marked and unmarked regions; here we have instead a partition into marked, unmarked and entangling regions[1]. It is established here that if the entangling region is much smaller than the marked region, then the search algorithm will still succeed with a high probability.

6 Conclusion and further work

Taking, as a performance baseline, the pure random search algorithm [3] (simply, randomly sampling the domain repeatedly), the multistart method provides a large improvement, for a wide class of optimisation problems. Separately, implementing pure adaptive search [13, 14] via Grover's algorithm [1, 4, 7, 9] provides an acceleration on the order of the square root of

---

[1]The good, the bad and the ugly.

the domain size; this is impressive for a method not relying on any problem structure, but quickly lags behind multistart for moderately structured problems. In [5] it was shown how the quadratic acceleration due to Grover's algorithm, and the acceleration (in proportion to the number of domain points in the deepest basin) due to multistart, could be achieved simultaneously.

It is now known that the quantum complexity of gradient estimation is smaller than its classical complexity by a factor which is linear in domain dimension [6, 10]. The question naturally arises whether we can simultaneously claim the benefit of this linear acceleration, alongside the accelerations due to Grover's algorithm and to multistart. This paper provides the answer. Section 2 describes the quantum basin hopper algorithm. Sections 3 to 5 study the error introduced into the algorithm by the indeterminacy inherent in quantum gradient estimation, using a "divide-and-conquer" approach. Section 3 splits the error in gradient estimation into "large" and "small" errors. Section 4 shows that "large" errors are unproblematic due to their low amplitude. Section 5 shows that "small" errors only corrupt Grover's algorithm on a small subset of the search space, with only slight overall effect.

To pursue this work further, a natural first step would seem to be to tighten the bounds established, in order to predict, for a given optimisation problem, what algorithm parameters and computational precision should be used. A careful analysis of the performance of a quantum-gradient-estimate implementation of, say, a conjugate gradient method on a quadratic form might yield insights.

References

[1] BARITOMPA, W. P., BULGER, D. W., and WOOD, G. R., *Grover's quantum algorithm applied to global optimisation*, SIAM Journal of Optimization, to appear.

[2] BOYER, M., BRASSARD, G., HØYER, P., and TAPP, A., *Tight bounds on quantum searching*, Fortschritte der Physik, Vol. 46, pp. 493–506, 1998.

[3] BROOKS, S. H., *A discussion of random methods for seeking maxima*, Operations Research, Vol. 6, pp. 244–251, 1958.

[4] BULGER, D. W., BARITOMPA, W. P., and WOOD, G. R., *Implementing pure adaptive search with Grover's quantum algorithm*, Journal of Optimization Theory and Applications, Vol. 116, pp. 517–529, 2003.

[5] BULGER, D. W., *Combining a local search and Grover's algorithm in black-box global optimisation*, preprint, `http://www.stat.mq.edu.au/research/papers/LocalGrover.pdf`, 2004.

[6] BULGER, D., *Quantum computational gradient estimation*, `http://lanl.arxiv.org/abs/quant-ph/0507109`, 13 Jul 2005.

[7] DÜRR, C., and HØYER, P., *A quantum algorithm for finding the minimum*, `http://lanl.arxiv.org/abs/quant-ph/9607014`, version 2, 7 Jan 1999.

[8] GOĆWIN, M., *Complexity of searching maximum of a function on quantum computer*, `http://lanl.arxiv.org/abs/quant-ph/0507060`, 6 Jul 2005.

[9] GROVER, L. K., *A fast quantum mechanical algorithm for database search*, Proceedings of the 28th Annual ACM Symposium on Theory of Computing, 1996.

[10] JORDAN, S. P., *Fast quantum algorithm for numerical gradient estimation*, Physical Review Letters, to appear (also `http://lanl.arxiv.org/abs/quant-ph/0405146`, version 2, 2 Jan 2005)

[11] NAYAK, A., and WU, F., *The quantum query complexity of approximating the median and related statistics*, Proceedings of the 31st Annual ACM Symposium on Theory of Computing, 1999.

[12] PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., and FLANNERY, B. P., *Numerical Recipes in C, second edition*, Cambridge University Press, 1995.

[13] ZABINSKY, Z. B., and SMITH, R. L., *Pure adaptive search in global optimization*, Mathematical Programming, Vol. 53, pp. 323–338, 1992.

[14] ZABINSKY, Z. B., WOOD, G. R., STEEL, M. A., and BARITOMPA, W. P., *Pure adaptive search for finite global optimization*, Mathematical Programming, Vol. 69, pp. 443–448, 1995.